1. Towards Upper-Triangulation By An Orthonormal Basis

Previously in this course, we have seen the value of changing our coordinates to be eigenbasis-aligned, because we can then view the system as a cascade of scalar systems. If we have a diagonalization, then these scalar equations are fully uncoupled, and can therefore be treated completely separately. But even when we cannot diagonalize, we can upper-triangularize in a way that allows us to solve the equations one at a time, from the "bottom up".

In this problem, to better understand the steps involved, we will use the following concrete example:

\[
S_{[3 \times 3]} = \begin{bmatrix}
\frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\
\frac{5}{12} & \frac{5}{12} & \frac{1}{6} \\
\frac{1}{6} & \frac{1}{6} & \frac{1}{3}
\end{bmatrix}
\]

(1)

and figure out the general case by abstracting variables. Note that there is a datahub link to a jupyter notebook on the website, which will allow you to perform the numerical calculations quickly to connect the symbolic analysis to an example, but without being time-consuming in the process.¹

(a) Consider a non-zero vector \( \vec{u}_0 \in \mathbb{R}^n \). Can you think of a way to extend it to a set of basis vectors for \( \mathbb{R}^n \)? In other words, find \( \vec{u}_1, \cdots, \vec{u}_{n-1} \), such that \( \text{span}(\vec{u}_0, \vec{u}_1, \cdots, \vec{u}_{n-1}) = \mathbb{R}^n \). To begin with, consider

\[
\begin{bmatrix}
1 \\
-1 \\
0
\end{bmatrix}
\]

Can you get an orthonormal basis from what you just constructed?

*Hint: what was the last discussion all about? Also, the given vector isn’t normalized yet!*

¹This particular matrix has an additional special property of symmetry, but we won’t be invoking that here.

\[
q_1 = \begin{bmatrix}
\frac{\sqrt{2}}{2} \\
\frac{\sqrt{2}}{2} \\
0
\end{bmatrix},
q_2 = \begin{bmatrix}
0 \\
0 \\
1
\end{bmatrix},
q_3 = \begin{bmatrix}
0 \\
1 \\
0
\end{bmatrix}
\]

\[
\begin{bmatrix}
\frac{1}{\sqrt{2}} \\
\frac{1}{\sqrt{2}} \\
0
\end{bmatrix} 
\Rightarrow 
q_2' = \begin{bmatrix}
\frac{\sqrt{2}}{2} \\
\frac{-\sqrt{2}}{2} \\
0
\end{bmatrix}
\]

\[
Q_z = \begin{bmatrix}
\frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\
\frac{-\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} \\
0 & 1 & 0
\end{bmatrix}
\]
(b) Now consider a real eigenvalue $\lambda_1$, and the corresponding (normalized) eigenvector $\bar{v}_1 \in \mathbb{R}^n$ of a square matrix $M \in \mathbb{R}^{n \times n}$. From the previous part, we know that we can extend $\bar{v}_1$ to an orthonormal basis of $\mathbb{R}^n$. We will denote the basis by

$$U = \begin{bmatrix}
\bar{u}_1 & \bar{u}_2 & \ldots & \bar{u}_n
\end{bmatrix}$$

where $\bar{u}_1 = \bar{v}_1$ (note that this eigenvector is already normalized).

Our goal is to look at what the matrix $M$ looks like in the coordinate system defined by the basis $U$.

Compute $U^T MU$ by writing $U = \begin{bmatrix} \bar{v}_1 & R \end{bmatrix}$, where $R \triangleq \begin{bmatrix} \bar{r}_1 & \bar{r}_2 & \ldots & \bar{r}_{n-1} \end{bmatrix}$.

$$U^T MU = \begin{bmatrix} v_1^T & v_n^T \end{bmatrix} M \begin{bmatrix} v_1 & R \end{bmatrix} = \begin{bmatrix} v_1^T & v_n^T \end{bmatrix} \begin{bmatrix} M v_1 & MR \\
R v_n & RMR \end{bmatrix}$$

$$= \begin{bmatrix} v_1^T M v_1 & v_1^T MR \\
R v_n^T M v_1 & R v_n^T MR \end{bmatrix} = \begin{bmatrix} \lambda_1 & v_1^T MR \\
0 & R v_n^T MR \end{bmatrix}$$

$$v_1^T M v_1 = v_1^T \lambda_1 = \lambda_1$$

$$R v_n^T M v_1 = R R v_n^T \lambda_1 = 0$$

(c) Show that $U^{-1} = U^T$.

$$U^T U = I$$

$$U^T U = \begin{bmatrix} u_1^T & u_2^T & \ldots & u_n^T \end{bmatrix} \begin{bmatrix} 1 & & & \\
& 1 & & \\
& & \ddots & \\
& & & 1 \end{bmatrix} = \begin{bmatrix} u_1^T u_1 & & & \\
& u_2^T u_2 & & \\
& & \ddots & \\
& & & u_n^T u_n \end{bmatrix} = I$$

if $i = j$, $u_i^T u_j = 1$

if $i \neq j$, $u_i^T u_j = 0$

$U^T U = I \Rightarrow U^T = U^{-1}$
(d) Define $Q = R^\top MR$. Look at the first column and the first row of $U^\top MU$ and show that:

$$M = U \begin{bmatrix} \lambda_1 & \bar{a}^\top \\ 0 & Q \end{bmatrix} U^\top$$

Here, $\bar{a}$ is a symbolic vector related to $M$, $R$, and $\bar{v}_1$ (we will show the relation!).

$$U^\top MU = \begin{bmatrix} \lambda_1 & \bar{v}_1^\top MR \\ 0 & R^\top MR \end{bmatrix}$$

$$U^\top M = \begin{bmatrix} \lambda_1 & \bar{v}_1^\top MR \\ 0 & R^\top MR \end{bmatrix} U^\top$$

$$M = U \begin{bmatrix} \lambda_1 & \bar{v}_1^\top MR \\ 0 & R^\top MR \end{bmatrix} U^\top$$
(e) Now, we can recurse on $Q$ to get:

$$Q = \begin{bmatrix} \mathbf{v}_2 & \mathbf{Y} \end{bmatrix} \begin{bmatrix} \lambda_2 & \mathbf{b}^T \\ 0 & P \end{bmatrix} \begin{bmatrix} \mathbf{v}_2 & \mathbf{Y} \end{bmatrix}^T$$

where we have taken $\mathbf{v}_2 \in \mathbb{R}^{n-1}$, a normalized eigenvector of $Q$, associated with eigenvalue $\lambda_2$. Again $\mathbf{v}_2$ is extended into an orthonormal basis to form $[\mathbf{v}_2 \ \mathbf{Y}]$.

Plug this form of $Q$ into $M$ above, to show that:

$$M = \begin{bmatrix} \mathbf{v}_1 & \mathbf{R} \mathbf{v}_2 & \mathbf{R} \mathbf{Y} \end{bmatrix} \begin{bmatrix} \lambda_1 & \tilde{\mathbf{a}}_1 & \tilde{\mathbf{a}}_{\text{rest}}^T \\ 0 & \lambda_2 & \mathbf{b}^T \\ 0 & 0 & P \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 & \mathbf{R} \mathbf{v}_2 & \mathbf{R} \mathbf{Y} \end{bmatrix}^T$$

where we define $\tilde{\mathbf{a}}$ to be the "adjusted" $\mathbf{a}$ to account for the substitution of $Q$; $\tilde{\mathbf{a}}^T = \mathbf{a}^T [\mathbf{v}_2 \ \mathbf{Y}]$.
• Neelesh Ramachandran.
• Yuxun Zhou.
• Edward Wang.
• Anant Sahai.
• Sanjit Batra.
• Pavan Bhargava.